

Original Article

Correlation surprise

Received (in revised form): 12th December 2013

Will Kinlaw

is Senior Managing Director and Head of Portfolio and Risk Management Research at State Street Associates and has been awarded the CFA designation.

David Turkington

is Managing Director in the Portfolio and Risk Management Research group at State Street Associates and has been awarded the CFA designation.

Correspondence: David Turkington, State Street Associates, 140 Mt. Auburn Street, Cambridge, Massachusetts 02138, USA

ABSTRACT Soon after Harry Markowitz published his landmark 1952 article on portfolio selection, the correlation coefficient assumed vital significance as a measure of diversification and an input to portfolio construction. However, investors typically overlook the potential for correlation patterns to help predict subsequent return and risk. Kritzman and Li (2010) introduced what is perhaps the first measure to capture the degree of multivariate asset price ‘unusualness’ through time. Their financial turbulence score spikes when asset prices ‘behave in an uncharacteristic fashion, including extreme price moves, decoupling of correlated assets, and convergence of uncorrelated assets.’ We extend Kritzman and Li’s study by disentangling the volatility and correlation components of turbulence to derive a measure of correlation surprise. We show how correlation surprise is orthogonal to volatility and present empirical evidence that it contains incremental forward-looking information. On average, after controlling for volatility, we find that periods characterized by correlation surprise lead to higher risk and lower returns to risk premia than periods characterized by typical correlations. This result holds across many markets including US equities, European equities and foreign exchange. Our results corroborate the predictive capacity of turbulence and suggest that its decomposition may also prove fruitful in forecasting investment performance.

Journal of Asset Management (2014) **14**, 385–399. doi:10.1057/jam.2013.27;
published online 9 January 2014

Keywords: correlation; turbulence; volatility; dislocation; risk management

The online version of this article is available Open Access

INTRODUCTION

Soon after Harry Markowitz published his landmark 1952 article on portfolio selection, the correlation coefficient assumed vital significance as a measure of diversification and an input to portfolio construction. More recently, investors have come to recognize the importance of correlation to a wide variety of investment activities. Analysts have used the parameter to detect regime shifts, describe

markets as ‘risk-on/risk-off’ and justify the underperformance of stock pickers. Of course, to monitor the tangled web of relationships between assets can be daunting. To cover a universe of 10 assets, one must track 45 pair-wise correlations. Kritzman and Li (2010) introduced what is perhaps the first measure to capture the degree of correlation ‘unusualness’ across a set of assets through time. Their financial turbulence score spikes when asset

prices ‘behave in an uncharacteristic fashion, including extreme price moves, decoupling of correlated assets, and convergence of uncorrelated assets’. For at least two reasons, this framework is well suited to the purpose of quantifying correlation surprises. First, it summarizes in a single measure the collective unusualness of correlations across any universe of assets. Second, rather than identify whether correlations are high or low, it measures the degree to which interactions depart from their historical norms, whatever those may be.

In this article, we extend Kritzman and Li’s study by disentangling the volatility and correlation components of turbulence to derive a measure of correlation surprise. We also show how correlation surprise is orthogonal to volatility and present empirical evidence that it contains incremental forward-looking information. On average, after controlling for volatility, we find that periods characterized by correlation surprise lead to higher risk and lower returns to risk premia than periods characterized by typical correlations. This result holds across many markets including US equities, European equities and foreign exchange. Our results corroborate the predictive capacity of turbulence and suggest that its decomposition may also prove fruitful in forecasting investment performance.

This article is organized as follows. First, we review the methodology behind Kritzman and Li’s financial turbulence measure, also known as the Mahalanobis distance, and review its empirical features. Next, we show how to decompose turbulence to isolate the contribution of correlation surprises. We then present empirical evidence that correlation surprises contain incremental information about future risk and return at both daily and monthly frequencies.

THE MAHALANOBIS DISTANCE AS A MEASURE OF FINANCIAL TURBULENCE

Using a formula originally developed by Mahalanobis (1927, 1936) to categorize human

skulls and later employed by Chow *et al* (1999) to stress test portfolios for turbulent periods, Kritzman and Li (2010) define their financial turbulence statistic as a multivariate unusualness score. We adopt the same definition of turbulence but simply divide by the number of assets (which is constant through time) in order to facilitate ease of interpretation in subsequent sections of this article:

$$d_t = (\gamma_t - \mu) \Sigma^{-1} (\gamma_t - \mu) / n \quad (1)$$

where

d_t	turbulence for a particular time period t (scalar)
γ_t	vector of asset returns for period t ($1 \times n$ vector)
μ	sample average vector of historical returns ($1 \times n$ vector)
Σ	sample covariance matrix of historical returns ($n \times n$ matrix)
n	number of assets in universe

This statistic, which can be thought of as a multivariate z -score, measures the statistical unusualness of a contemporaneous cross-section of asset returns relative to its historical distribution. It captures the extent to which the risk-adjusted magnitudes of the returns differ from their historical means as well as the extent to which their interaction is inconsistent with the historical correlation matrix. Turbulence is different from cross-sectional volatility, which measures the dispersion around the cross-sectional mean but ignores the time series means.¹ It is also different from the rolling volatility of a portfolio because it describes the unusualness of a particular day rather than the dispersion in returns over a period of time.

Empirically, turbulence has several attractive features:

- It tends to spike during recognizable periods of market stress that are characterized by heightened volatility and correlation breakdowns.

- It is linked to investment performance; on average, returns to a wide variety risk premia are significantly lower during turbulent periods.²
- It is persistent. Turbulent episodes tend to cluster in time and do not subside immediately after they arise.

Taken together, the persistence of turbulence and its link with returns suggest that investors could enhance performance by de-risking when turbulence first strikes. For example, Kritzman and Li show how to improve the performance of a foreign exchange carry strategy by reducing exposure when currency turbulence rises. In this article, we put the persistence of turbulence – and its relationship with subsequent return and risk – under the microscope by disentangling correlation surprises from magnitude surprises.

ISOLATING CORRELATION SURPRISES

The turbulence methodology is ideally suited to detecting periods when the co-movement between assets differs from what we would expect based on historical correlations. For a given day, the turbulence score captures both the average degree of unusualness in individual asset returns (magnitude surprise) and the degree of unusualness in the interaction between each pair of assets (correlation surprise). To disentangle these two components and isolate the degree of correlation surprise, we first compute the magnitude surprise. Magnitude surprise is equal to the turbulence score, given in equation (1), where all off-diagonal elements in the covariance matrix are set to zero.³ This ‘correlation-blind’ turbulence measure captures magnitude surprises, but ignores whether co-movement is typical or atypical. Next, we divide the standard turbulence score – which includes correlation effects – by the magnitude surprise. This ratio is the correlation surprise: the unusualness of interactions on a particular day relative to history.

A correlation surprise ratio greater than one is associated with correlation breakdowns (that is, previously correlated assets diverging and/or previously negatively correlated assets converging). A correlation surprise ratio less than one is associated with relatively typical correlation outcomes. In other words, days characterized by low correlation surprise are actually less unusual than the magnitudes of the individual returns alone would suggest. To review, we compute the following quantities to calculate correlation surprise:

1. Magnitude surprise: a ‘correlation-blind’ turbulence score in which all off-diagonals in the covariance matrix are set to zero.
2. Turbulence score: the degree of statistical unusualness across assets on a given day, as given in equation (1).
3. Correlation surprise: the ratio of turbulence to magnitude surprise, using the above quantities (2) and (1), respectively.

$$\text{Correlation surprise} = \frac{\text{Turbulence}}{\text{Magnitude surprise}} \quad (2)$$

To build intuition around the turbulence and correlation surprise scores, let us start by considering a single asset, x . In this case, turbulence is simply equal to the squared z -score of the asset return, as shown in equation (3). To simplify the formulas in this section, we will assume (without loss of generality) that the average value of all relevant return streams are zero, and we will simply denote the asset's current period return observation as x .

$$\begin{aligned} \text{Turbulence for one asset} &= x(\sigma_x^2)^{-1}x \\ &= \left(\frac{x}{\sigma_x}\right)^2 = z_x^2 \quad (3) \end{aligned}$$

It is impossible, by definition, for one asset to exhibit any correlation surprise. If we consider two assets, we can think of turbulence as a multivariate z -score.

Rather than normalizing only by the variance of an asset, as in equation (3), turbulence now normalizes for the entire covariance matrix of the assets, which accounts for their historical variances and the correlation between them as shown in equation (4).

Turbulence for two assets

$$= (x \quad y) \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \quad (4)$$

By dividing turbulence by magnitude surprise, we can express the correlation surprise for two assets as shown in equation (5) (the derivation of this formula is provided in the Appendix). We see that correlation surprise is expressed in terms of normalized asset z -scores and the historical correlation, ρ . All units of magnitude cancel out in this formula.

Correlation surprise for two assets

$$= \frac{1}{1-\rho^2} \cdot \left(1 - \frac{\rho z_x z_y}{\frac{1}{2}(z_x^2 + z_y^2)} \right) \quad (5)$$

Turbulence and correlation surprise also have intuitive geometric interpretations. Consider a simple empirical example in which our universe consists of two assets, A and B, with means of zero, volatilities of 5 per cent, and a correlation of 0.5. Figure 1 shows the turbulence score, magnitude surprise and correlation surprise for two multivariate return observations.

The dashed ellipse in Figure 1 is the iso-turbulence ellipse. All observations that fall along this ellipse have the same turbulence score. Its slant reflects the positive correlation between the two assets: in any given period, it is more likely that A and B move in the same direction than in opposite directions. In other words, when A and B move in the same direction, we require a larger return magnitude to produce the same degree of turbulence. In this example, period 2 is more turbulent than period 1 despite the fact that the magnitudes of the two

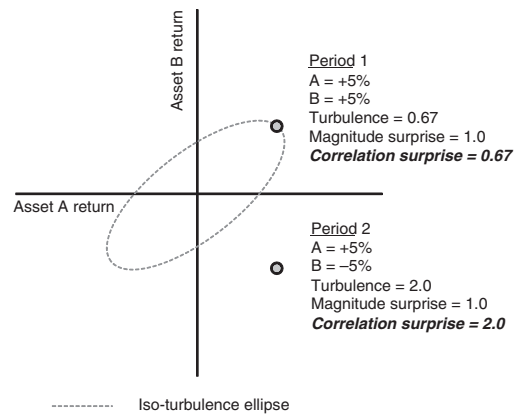


Figure 1: The iso-turbulence ellipse and two return observations for assets A and B.

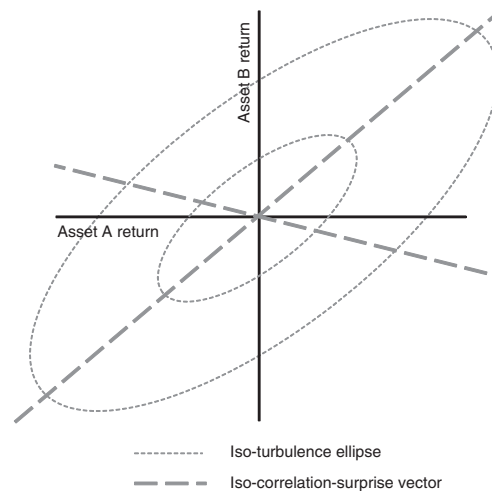


Figure 2: Iso-correlation surprise vectors.

observations are identical (each has a magnitude surprise of 1.0). Period 2 has a higher correlation surprise than period 1 because it reflects an outcome where the two assets, which are expected to move together, diverge.

The correlation surprise score is akin to a compass: it contains information about the multivariate direction along which the observation falls. It does not measure the magnitude of returns. The diagonal lines in Figure 2 are iso-correlation surprise vectors; regardless of their magnitude, any two points along one of these lines will have the same correlation surprise score.

HOW DOES CORRELATION SURPRISE DIFFER FROM OTHER MEASURES THAT INCORPORATE CORRELATION?

To our knowledge, correlation surprise is the only measure that isolates the degree to which the co-movement across a set of assets is typical or atypical relative to history.⁴ Below we list three other measures that are sensitive to correlation shifts and describe the important ways in which they fail to capture correlation surprises.

- *Rolling correlation:* Investors sometimes monitor the rolling correlation between two assets to help them identify regime shifts. Whereas rolling correlation is a pair-wise measure, correlation surprise can capture in a single parameter the unusualness of co-movement across a large universe of assets.⁵ Indeed, the number of pair-wise correlations quickly becomes unmanageable as the asset universe expands. Again, to monitor a 10-asset universe, we would be required to estimate 45 rolling correlations; for a 100-asset universe, the number increases to 4950.⁶
- *Cross-sectional volatility:* Cross-sectional volatility is different from correlation surprise because it fails to account for the degree to which a particular correlation outcome is typical or atypical. Put differently, cross-sectional volatility spikes when asset returns diverge, regardless of whether their typical correlation is 10 per cent or 90 per cent. All else equal, a divergence is far more unusual in the latter case than in the former.
- *Multivariate Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models:* This class of models extends the univariate GARCH framework to derive volatility forecasts based on lagged covariance terms as well as lagged variance terms. For example, a multivariate GARCH model might forecast the

volatility of a two-stock portfolio based on lagged innovations in stock A, lagged innovations in stock B and lagged covariance terms between stocks A and B.⁷ Engle (2002) has proposed a Dynamic Conditional Correlation model, which is a simple class of multivariate GARCH models. However, to our knowledge, none of the myriad multivariate GARCH specifications account for interactions between variance and covariance terms.⁸ And, as we will demonstrate in the next section, volatile episodes characterized by atypical correlations tend to be more persistent and severe than volatile episodes characterized by typical correlations. The decomposition of the Mahalanobis distance enables us to analyze the intertemporal relationship between correlation and magnitude surprises in a way that multivariate GARCH models do not.

DATA AND RESULTS

We construct correlation surprise series for three asset universes: US equities, European equities and currencies. Table 1 provides details regarding the component indices, start date and lookback window for each series.

Figure 3 shows a scatter-plot of daily correlation surprise versus magnitude surprise scores for each asset universe. For the purposes of this figure, we converted each metric into a per cent rank with regard to its full sample distribution. The dark lines show the best fit linear regression line for each plot.

It is apparent from visual inspection of Figure 3 that the correlation and magnitude surprise scores capture different information (which is collectively captured in the turbulence score). In fact, on a contemporaneous basis the two metrics are negatively correlated. What is intriguing is that despite this contemporaneous negative correlation, we find strong empirical evidence that high correlation surprise tends to precede

Table 1: Time series data

	<i>US equities</i>	<i>European equities</i>	<i>Currencies</i>
Lookback window	10 years of daily returns ^a	10 years of daily returns ^a	3 years of daily returns
Index start date	26 November 1975	26 November 1975	24 November 1977
Index end date	30 September 2010	30 September 2010	30 September 2010
Data source	S&P US Sectors ^b	MSCI Europe Sectors ^b	WMR 4 pm London Fix Rates vs the US dollar
Constituents used to build index	Consumer discretionary Consumer staples Energy Financials Healthcare Industrials Information technology Materials Telecommunications Utilities	Consumer discretionary Consumer staples Energy Financials Healthcare Industrials Information technology Materials Telecommunications Utilities	Australian dollar British pound Canadian dollar Euro ^c Japanese yen New Zealand dollar Norwegian krone Swedish krona Swiss franc —

^aReturns are equally weighted to calculate mean and covariance. To capture longer cycles in the equity markets, we begin with a window of 3 years which is grown to 10 years and rolled forward.

^bDatastream sector data is used prior to 1995, when MSCI daily data becomes unavailable. The S&P 500[®] index and its GICS[®] level 1 sector sub-indices are proprietary to and are calculated, distributed and marketed by S&P Opco, LLC (a subsidiary of S&P Dow Jones Indices LLC) ('SPDJ'), its affiliates and/or its licensors and have been licensed for use. S&P[®], S&P 500[®] and GICS[®], among other famous marks, are registered trademarks of Standard & Poor's Financial Services LLC, and Dow Jones[®] is a registered trademark of Dow Jones Trademark Holdings LLC. Neither SPDJI, its affiliates nor their third-party licensors make any representation or warranty, express or implied, as to the ability of any index to accurately represent the asset class or market sector that it purports to represent. Neither SPDJI, its affiliates nor their third-party licensors shall have any liability for any errors or omissions related to its indices or the data included therein nor do the views and opinions of the authors expressed herein necessarily state or reflect those of SPDJI, its affiliates and/or its licensors. SPDJI, its affiliates and their licensors reserve all rights with respect to the S&P 500[®] index and its GICS[®] level 1 sector sub-indices.

^cDeutsche mark used prior to the introduction of the euro.

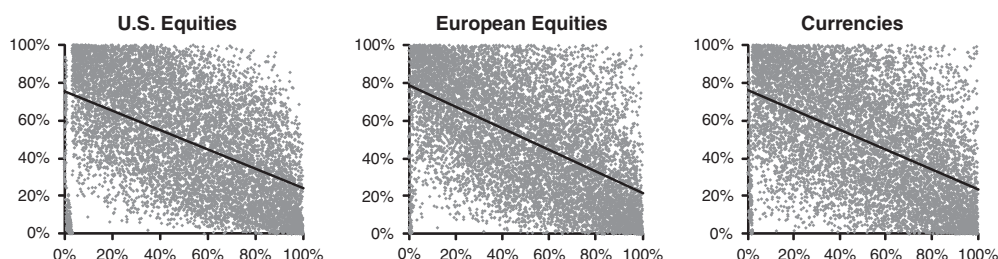


Figure 3: Daily correlation surprise (vertical axis) versus magnitude surprise (horizontal axis).

higher risk and also lower returns. We will discuss these findings in more detail shortly. Before we shift our focus to empirics, let us first consider a more theoretical question: why would we expect unusual correlations to precede heightened volatility and lower returns? There are several plausible reasons. Investors who build correlation assumptions into their models – either explicitly or through intuition – may underperform when correlations deviate from their historical

norms, inducing them to de-risk. In addition, financial markets are not perfectly efficient and it takes some measure of time for information to propagate from one segment to another. For example, consumer stocks may react immediately to a particular news item, but financial stocks may not react until investors have analyzed relationships between these sectors, many of which are obscure and opaque. In this scenario, a shock would register as a correlation event before it

registers as a market-wide volatility event. It is also possible that there is a behavioral explanation. Perhaps investors tend to de-risk when markets are ‘acting weird’ and are difficult to understand.

Whatever the reason, we find convincing empirical evidence that there is a lead-lag relationship between correlation surprises, volatility and returns. For a vivid example of how correlation surprise manifests in practice, consider the month of September 2008, one of the most turbulent months in financial history. A correlation surprise score above one occurred on nine days that month.⁹ The average magnitude surprise score on the days following these correlation surprise events was 8.7, and the average *daily* S&P 500 return was –219 basis points (bps). In contrast, the days with correlation surprise less than one were followed by an average magnitude surprise of 4.0 and an average S&P 500 return of +86 bps. The worst one-day market loss during the month occurred on 29 September, when the S&P 500 lost 879 bps from the previous day’s close, accompanied by a very large magnitude surprise of 33. Correlation surprise on the day of this drawdown was low (0.4), but on the previous day it was very high (1.9), in part because of a large divergence in the daily return of the Materials sector (–252 bps) and the Financials Sector (+309 bps), which over the preceding 10 years had experienced a positive correlation of 57 per cent.¹⁰ A similar divergence occurred on Friday, 12 September, with correlation surprise of 1.7 reflecting a gain for Materials (+323 bps) and a loss for Financials (–106 bps). The market dropped 471 bps on Monday, following news of Lehman’s default.

Monday had low correlation surprise, and the market rallied 175 bps on Tuesday. However, Tuesday’s correlation surprise was 2.0.

The market dropped another 471 bps on Wednesday. This example is clearly anecdotal and in choosing it we are guilty of selection bias. However, we find that the same intuition holds on average across three different markets from the 1970s through 2010. We will now turn to a more robust empirical analysis.

For each of our three universes, we perform a series of experiments consisting of three broad steps:

1. Identify the 20 per cent of days in the historical sample with the highest magnitude surprise scores.¹¹
2. Partition the sample from step 1 into two smaller subsamples: days with high correlation surprise (greater than one) and days with low correlation surprise (less than one).
3. Measure, for the full sample identified in step 1 and its two subsamples identified in step 2, the subsequent volatility and performance of relevant investments and strategies.

Before presenting out-of-sample results, we examine the contemporaneous relationship between magnitude surprise and correlation surprise. Table 2 shows the average magnitude surprise for each of the three subsamples described in steps 1 through 3, above. Of the 20 per cent most ‘volatile’ days in each universe (as determined by magnitude surprise), the days characterized by high correlation surprise exhibit less magnitude surprise on average than the days

Table 2: Conditional average magnitude surprise on the *day of the reading*

	<i>US equities</i>	<i>European equities</i>	<i>Currencies</i>
Day of top 20% MS with CS ≤ 1	4.6	4.0	3.6
Day of top 20% MS (all observations)	3.9	4.0	3.5
Day of top 20% MS with CS > 1	2.6	2.8	3.0
Difference in means (high CS minus low CS sample)	–2.0	–1.2	–0.6
Percentage increase (high CS vs low CS sample)	–43%	–29%	–17%

Table 3: Conditional average magnitude surprise on the *day after the reading*

	US Equities	European Equities	Currencies
Day following top 20% MS with $CS \leq 1$	2.1	2.1	1.5
Day following top 20% MS (all observations)	2.3	2.4	1.7
Day following top 20% MS with $CS > 1$	2.6	3.0	2.1
Difference in means (high CS minus low CS sample)	0.5	0.9	0.6
Percentage increase (high CS vs low CS sample)	21%	41%	38%
<i>t</i> -statistic of difference in means test	1.54	3.12	3.02
<i>p</i> -value of difference in means test	0.06	0.00	0.00

characterized by low correlation surprise. In other words, the days characterized by the most extreme volatility tended to exhibit more 'typical' correlations.

Table 3 shows the same results for the day after the reading. Now, the pattern is reversed: for each universe, the days characterized by the most extreme magnitude surprise tended to be preceded by atypical correlations. In other words, heightened volatility coupled with unusual correlations foretells higher next-day volatility than heightened volatility coupled with typical correlations. The *t*-statistics and *p*-values in Table 3 reveal that these differences are statistically significant at the 90 per cent level.

These results suggest that correlation surprises contain incremental information about future volatility. However, most investors are more interested in future returns. To evaluate the relationship between correlation surprise and subsequent returns, we employ the same three-step test. But instead of measuring magnitude surprise, we measure return, standard deviation and hit rate for three investable indices. (Hit rate is defined as the percentage of days following the given signal that experience positive returns.) For US and European equities, we analyze the return of the corresponding index listed in Table 1. For currencies, we analyze the returns of a simple carry strategy.¹² Table 4 presents these results, along with the number of observations (days) in each subsample.

For all three investable indices, Table 4 shows that the average return is lower following extreme magnitude days characterized by high correlation surprise

than following extreme magnitude days characterized by low correlation surprise.

These results are statistically significant at the 95 per cent level for US equities.

Interestingly, magnitude surprise alone is not particularly effective for partitioning the next day's returns. In fact, for US equities, the 20 per cent of days with the largest magnitude surprise scores actually foretold higher returns, on average, than the remaining 80 per cent of the sample. Table 4 also reveals that the hit rate (per cent of positive days) for all three indices is lowest following days characterized by both high correlation surprise and high magnitude surprise.¹³ Finally, Table 4 shows that the standard deviation of returns is higher following days where both magnitude surprise and correlation surprise are high.¹⁴ This result holds for all three asset classes.

To further explore the relationship between correlation surprise and volatility, we analyze the performance of a hypothetical short volatility strategy. Specifically, we simulate selling an at-the-money straddle: simultaneously writing an at-the-money put option and an at-the-money call option. If the price of the reference asset remains unchanged, then the options expire worthless and the seller pockets the premium. But this strategy can suffer spectacular losses when volatility spikes and one of the options expires in the money. Table 5 shows next-day performance of short volatility strategies in the US equity, European equity and currency markets. Strictly speaking, these short volatility proxies are not investable because their performance is approximated using the Black-Scholes-Merton model (based on

Table 4: Investment performance on the *day after* the reading^a

		Average (annualized)	Std. deviation (annualized) (%)	Hit rate (% days positive)	Number of days in sample
US equities, S&P 500 ^b	Full sample	7.2%	17.1	51.0	9092
	Day following top 20% MS with CS ≤1	27.7%	23.8	57.3	1211
	Day following top 20% MS (all observations)	16.2%	24.5	54.3	1818
	Day following top 20% MS with CS >1	-6.7%	25.9	48.1	607
	Difference (high CS minus low CS sample)	-34.5%	2.2	-9.2	—
	<i>t</i> -statistic of difference in means test	-1.73	—	—	—
	<i>p</i> -value of difference in means test	0.04	—	—	—
European equities, MSCI Europe ^b	Full sample	11.5%	16.1	55.0	9092
	Day following top 20% MS with CS ≤1	7.0%	21.2	54.0	1297
	Day following top 20% MS (all observations)	5.8%	22.9	53.3	1819
	Day following top 20% MS with CS >1	2.5%	26.7	51.5	522
	Difference (high CS minus low CS sample)	-4.5%	5.5	-2.5	—
	<i>t</i> -statistic of difference in means test	-0.22	—	—	—
	<i>p</i> -value of difference in means test	0.41	—	—	—
Currencies, G10 carry strategy ^b	Full sample	3.4%	6.1	54.9	5024
	Day following top 20% MS with CS ≤1	-1.2%	6.9	54.6	784
	Day following top 20% MS (all observations)	-3.1%	7.8	53.7	1063
	Day following top 20% MS with CS >1	-8.3%	10.0	51.3	279
	Difference (high CS minus low CS sample)	-7.0%	3.1	-3.3	—
	<i>t</i> -statistic of difference in means test	-0.69	—	—	—
	<i>p</i> -value of difference in means test	0.26	—	—	—

^aWe present annualized values for ease of interpretation. We conduct all statistical tests using daily data. All results are out of sample. However, we selected the threshold for the 20% of days with the largest magnitude surprise based on the full sample. Choosing other volatility thresholds yielded similar results.

^bS&P 500 and MSCI Europe results from Nov 1975 to Sep 2010, currencies from Oct 1990 to Dec 2009.

at-the-money implied volatility) as opposed to market prices for options. But they nonetheless provide insight into the relationship between correlation surprise and the short volatility premium.

Table 5 reveals that the average annualized return for all three strategies was lowest following the subsample characterized by high magnitude surprise and high correlation surprise simultaneously. The volatility was highest following this subsample. The hit rate (per cent of days with a positive return) was lowest following the joint occurrence in the currency market but highest in the two equity markets, where the signal appears to suffer from some false positives. Nonetheless, overall, these results suggest that a short

volatility investor would be well advised to hedge his or her exposure when magnitude surprise and correlation surprise spike simultaneously.

HOW QUICKLY DOES THE SIGNAL DECAY?

Thus far, our analysis has focused on the relationship between magnitude and correlation surprise on one day and performance on the next. An investor would need to react very quickly to exploit this information. Table 6 presents an analysis of how quickly a one-day spike in the signal decays. Specifically, for the experiments summarized in Tables 4 and 5, we report the

Table 5: Short straddle performance *following* conditioned and unconditioned spikes in magnitude surprise

		Average (annualized)	Std. deviation (annualized) (%)	Hit rate (% days positive)	Number of days in sample
US equities, Short S&P 500 straddle ^a	Full sample	13.3%	11.2	63.5	5371
	Day following top 20% MS with CS ≤ 1	31.8%	16.5	64.0	726
	Day following top 20% MS (all observations)	24.9%	17.0	64.4	1241
	Day following top 20% MS with CS > 1	15.1%	17.7	64.9	515
	Difference (high CS minus low CS sample)	-16.7%	1.2	0.8	—
	<i>t</i> -statistic of difference in means test	-1.06	—	—	—
	<i>p</i> -value of difference in means test	0.14	—	—	—
European equities, Short DAX 30 straddle ^a	Full sample	4.9%	14.1	59.8	4,869
	Day following top 20% MS with CS ≤ 1	15.1%	18.0	59.1	804
	Day following top 20% MS (all observations)	13.2%	20.8	59.2	1186
	Day following top 20% MS with CS > 1	9.1%	25.7	59.4	382
	Difference (high CS minus low CS sample)	-6.1%	7.7	0.3	—
	<i>t</i> -statistic of difference in means test	-0.26	—	—	—
	<i>p</i> -value of difference in means test	0.40	—	—	—
Currencies, Short a basket of G10 straddles vs USD ^a	Full sample	3.4%	6.1	54.9	3543
	Day following top 20% MS with CS ≤ 1	10.2%	6.4	59.5	603
	Day following top 20% MS (all observations)	6.2%	6.4	57.6	807
	Day following top 20% MS with CS > 1	-5.5%	6.6	52.0	204
	Difference (high CS minus low CS sample)	-15.6%	0.2	-7.5	—
	<i>t</i> -statistic of difference in means test	-1.86	—	—	—
	<i>p</i> -value of difference in means test	0.03	—	—	—

^aS&P 500 results from Mar 1990 to Sep 2010, DAX 30 results from Feb 1992 to Sep 2010, currency results from May 1996 to Dec 2009. Historical returns for at-the-money straddles are simulated assuming Black-Scholes-Merton pricing. Straddles are sold monthly for equities and weekly for currencies. The currency strategy represents an equally weighted basket of G10 short straddles versus the US dollar.

following values for the subsequent 20-day periods:

- average magnitude surprise for periods following high correlation surprise minus average magnitude surprise for periods following low correlation surprise, and
- average return during periods following high correlation surprise minus average return during periods following low correlation surprise.

We report average differentials for direct exposure to the equity indices and G10 carry strategy as well as to the simulated short-volatility strategies.

To better understand Table 6, consider the example of US equities. The average

magnitude surprise one day after a joint signal (a top-quintile magnitude surprise where correlation surprise was *greater* than one) was 0.6 higher than the average magnitude surprise one day after a magnitude-only signal (a top-quintile magnitude surprise where the correlation surprise was *less* than one). Over the next four days (days 2 through 5) this difference falls to 0.1. Table 6 also indicates that, based on our analysis spanning 1975 through 2010, the average return of US equities the day after a joint signal was 34.4 per cent lower than the day after a magnitude-only signal.

To summarize, we find a relationship between the joint signal and next-day volatility (as measured by magnitude surprise) across

Table 6: Time decay^a

After event, statistics for the following1 day	... days 2–5	...days 6–10	...days 11–20
US equities	Average magnitude surprise	0.6 (0.06)	0.1 (0.29)	0.1 (0.23)	0.2 (0.02)
	S&P 500 return (% , ann)	–34.5% (0.04)	4.4% (0.3)	–1.5% (0.42)	–0.9% (0.43)
	S&P 500 short straddle return (% , ann)	–16.7% (0.14)	–1.8% (0.39)	–1.7% (0.38)	–1.1% (0.38)
European equities	Average magnitude surprise	0.9 (0)	0.4 (0.02)	0.4 (0.01)	0.3 (0.02)
	MSCI Europe return (% , ann)	–4.5% (0.41)	–2.1% (0.41)	2.1% (0.4)	–8.6% (0.05)
	DAX 30 short straddle return (% , ann)	–6.1% (0.4)	–3.0% (0.38)	–5.8% (0.24)	0.8% (0.44)
Currencies	Average magnitude surprise	0.6 (0)	0.4 (0)	0.2 (0.01)	0.1 (0.03)
	G10 carry return (% , ann)	–7.0% (0.26)	–0.2% (0.49)	–5.3% (0.11)	–4.3% (0.05)
	G10 short straddles return (% , ann)	–15.6% (0.03)	–7.1% (0.07)	0.1% (0.49)	–0.8% (0.38)

p-values for difference-in-means tests shown in parentheses.

^aTime periods differ for each test. Please refer to the notes in Table 4 and Table 5 for specific dates.

all three markets. We also find a relationship between the joint signal and next-day returns. Of the three asset classes we studied, both of these relationships appear to decay most quickly in the US equity market. For the short volatility strategies, the results are most persistent in the currency market. In the next section, we explore a way to derive correlation surprise signals that apply to longer holding periods.

LONGER HORIZON SIGNALS

So far, we have viewed correlation surprise on a daily frequency. However, not all investors have the ability to react so quickly to changes in the market environment. Investors who have the ability to reallocate assets on a less frequent basis – such as monthly – can gain information from longer-term trends in magnitude surprise and correlation surprise. We create monthly signals by aggregating the daily time series within each month. For magnitude surprise, we equally weight all the daily observations within each month. For correlation surprise, however, we do not expect all observations to be equally informative. For example, if assets

only move by a few basis points on a given day, their pattern of co-movement may not reflect much more than random noise. But if assets move significantly, their pattern of co-movement is much more likely to represent an important fracture in the market. Therefore, we compute a weighted average of correlation surprise for each month using the daily magnitude surprise scores as weights. Mathematically, we calculate the following quantities each month, where *T* represents the total number of observations, *MS* is daily magnitude surprise and *CS* is daily correlation surprise:

$$MonthMS_t = \frac{1}{T} \sum_{i=1}^T MS_i$$

$$MonthCS_t = \frac{\sum_{i=1}^T CS_i MS_i}{\sum_{j=1}^T MS_j}$$

Across all asset classes in our study, the monthly correlation surprise series is negatively correlated to the moving average magnitude surprise series. Using these monthly signals, we repeat the conditional returns analysis from Tables 4 and 5 for

Table 7: Investment performance the *month following* conditioned and unconditioned signals

		Average (annualized)	Std. deviation (annualized) (%)	Hit rate (% months positive)	Number of observations
US equities, S&P 500	Full sample	10.3%	15.6	61.4	417
	Month following high MS with low CS	10.8%	23.7	65.3	49
	Month following high MS (all observations)	7.2%	22.3	59.8	87
	Month following high MS with high CS	2.5%	20.6	52.6	38
	Difference (high CS minus low CS sample)	-8.4%	-3.1	-12.7	—
	<i>t</i> -statistic of difference in means test	-0.51	—	—	—
	<i>p</i> -value of difference in means test	0.31	—	—	—
European equities, MSCI Europe	Full sample	10.0%	16.5	63.5	417
	Month following high MS with low CS	16.4%	19.9	65.2	46
	Month following high MS (all observations)	8.2%	20.6	60.0	85
	Month following high MS with high CS	-1.5%	21.4	53.8	39
	Difference (high CS minus low CS sample)	-17.9%	1.4	-11.4	—
	<i>t</i> -statistic of difference in means test	-1.15	—	—	—
	<i>p</i> -value of difference in means test	0.13	—	—	—
Currencies, G10 carry strategy	Full sample	3.6%	5.8	65.2	227
	Month following high MS with low CS	-2.5%	8.5	46.7	30
	Month following high MS (all observations)	-4.1%	8.0	45.7	46
	Month following high MS with high CS	-7.1%	7.3	43.8	16
	Difference (high CS minus low CS sample)	-4.5%	-1.2	-2.9	—
	<i>t</i> -statistic of difference in means test	-0.55	—	—	—
	<i>p</i> -value of difference in means test	0.29	—	—	—

subsequent one-month holding periods, assuming that positions would only be modified at the end of any given month. The monthly tests cover the same time period as the daily tests presented earlier, and we use the same simple parameters: an 80th percentile threshold for magnitude surprise and a threshold of one for correlation surprise. Tables 7 and 8 show these results for the broad investable indices and short straddles, respectively. The results from this analysis across all three asset universes show that on average, subsequent one-month asset performance is more negative when high magnitude surprise is accompanied by high correlation surprise than when it is accompanied by low correlation surprise. This result is fully consistent with our findings on daily data. As we might expect, many of the annualized return differentials are smaller for monthly data than for daily data,

but they are still quite meaningful. Across all tests in Tables 7 and 8, the hit rates (percentage of positive one-month returns) decrease following increases in correlation surprise.¹⁵

SUMMARY

We describe how to decompose the Mahalanobis distance (also known as financial turbulence) to derive a measure of correlation surprise across a set of assets. We show, both conceptually and empirically, that our correlation surprise measure is distinct from and incremental to magnitude surprise. Whereas magnitude surprise describes the extent to which asset returns are extreme *vis-à-vis* their historical distributions, correlation surprise describes the extent to which their interaction is unusual given the historical correlation matrix. Finally, we construct

Table 8: Short straddle performance the *month following* conditioned and unconditioned signals

		Average (annualized)	Std. deviation (annualized) (%)	Hit rate (% months positive)	Number of observations
US equities, Short S&P 500 straddle	Full sample	12.6%	8.7	73.6	246
	Month following high MS with low CS	13.1%	13.8	77.1	35
	Month following high MS (all observations)	11.3%	12.2	72.5	69
	Month following high MS with high CS	9.5%	10.5	67.6	34
	Difference (high CS minus low CS sample)	-3.6%	-3.3	-9.5	—
	<i>t</i> -statistic of difference in means test	-0.35	—	—	—
	<i>p</i> -value of difference in means test	0.36	—	—	—
European equities, Short DAX 30 straddle	Full sample	5.9%	12.7	62.8	223
	Month following high MS with low CS	30.6%	17.1	90.3	31
	Month following high MS (all observations)	16.4%	15.5	72.3	65
	Month following high MS with high CS	3.4%	13.1	55.9	34
	Difference (high CS minus low CS sample)	-27.1%	-4.0	-34.4	—
	<i>t</i> -statistic of difference in means test	-2.06	—	—	—
	<i>p</i> -value of difference in means test	0.02	—	—	—
Currencies, Short a basket of G10 straddles vs USD	Full sample	4.0%	4.3	60.6	160
	Month following high MS with low CS	8.0%	6.1	69.6	23
	Month following high MS (all observations)	7.4%	5.9	69.4	36
	Month following high MS with high CS	6.6%	5.6	69.2	13
	Difference (high CS minus low CS sample)	-1.4%	-0.5	-0.3	—
	<i>t</i> -statistic of difference in means test	-0.20	—	—	—
	<i>p</i> -value of difference in means test	0.42	—	—	—

correlation surprise series for three different asset classes and present evidence that the joint occurrence of high magnitude surprise and high correlation surprise foretells higher volatility and lower return than high magnitude surprise in isolation.

These findings have several concrete investment implications. A superior understanding of correlation surprises could enhance trade execution algorithms, which rely on correlation estimates to manage opportunity costs. In addition, risk managers could improve their volatility forecasts by incorporating correlation surprises into their models. And, perhaps most interestingly, portfolio managers may be able to enhance their performance by de-risking when they observe correlation surprises coupled with heightened volatility.

ACKNOWLEDGEMENTS

The authors thank Megan Czasonis, Mark Kritzman and Chirag Patel for helpful comments and assistance.

NOTES

1. Imagine a period where all assets experience a 50 per cent loss. Turbulence would spike but cross-sectional volatility would be zero.
2. Examples include the equity risk premium, small cap premium, growth minus value, the FX carry trade and hedge fund returns.
3. It is also equal to the average squared z -score across the assets in our universe.
4. A wide variety of conventional risk measures incorporate both volatility and correlation but differ from correlation surprise in obvious ways. For example, the rolling standard deviation of a portfolio does not necessarily capture correlation surprises among its components. Consider a portfolio that is allocated to two stocks with a historical correlation of 0.75. Imagine that on a particular day, one stock realizes a 20 per cent gain while the other experiences a 20 per cent loss. The returns offset one another perfectly, the portfolio's return for the day is 0 per cent, and its rolling volatility declines. In this case, volatility fails to capture an extremely unusual correlation outcome. Volatility-based metrics (such as value at risk) and volatility forecasting models (such as ARCH and GARCH) fail to capture correlation surprises in similar fashion.
5. Another difference between rolling correlation and correlation surprise is that the former is, as its name suggests,

a rolling measure whereas the latter is a normalized measure of unusualness over a discrete period. The relationship between rolling correlation and correlation surprise is analogous to the relationship between rolling volatility and a z -score.

6. A correlation matrix of n assets contains $((n \times n) - n)/2$ distinct correlation coefficients.
7. Typically, innovation terms are equal to squared residuals and covariance terms are equal to the product of residuals.
8. Multivariate GARCH models suffer from the 'curse of dimensionality' in that the number of covariance terms increases nonlinearly with the number of assets. The addition of a variance-covariance interaction term would render them even more unwieldy. Indeed, much of the existing multivariate GARCH literature is dedicated to adaptations that restrict the number of parameters. For a nice discussion on this subject, see Fengler and Herwartz (2008).
9. The days with a correlation surprise score above one were 2, 3, 5, 10, 12, 16, 19, 24 and 26 September.
10. The worst daily return following a correlation surprise reading less than one was 9 September's loss of 341 basis points.
11. We compute magnitude surprise series using the components and lookback windows listed in Table 3. We pre-condition the experiments on the 20 per cent of days with the largest magnitude surprise to isolate the correlation surprise signals that are likely to contain the most information; days characterized by high correlation surprise but miniscule returns are most likely noise.
12. To compute the carry returns, we assume equal sized long or short forward positions in each G10 currency (rebalanced monthly) depending on the currencies' interest rate differentials versus the US dollar.
13. The hit rate also provides information about false positives. For example, Table 4 indicates that, in the US equity market, we observed positive returns after a joint signal 48.1 per cent of the time.
14. It should be noted that the subsample means are significantly different from one another; hence, if we were to compute these volatilities around the full sample means we would expect to see even larger differences.
15. These results do not depend on month-end rebalancing. Although not reported, we obtained very similar results based on overlapping monthly holding periods.

REFERENCES

- Chow, G., Jacquier, E., Lowry, K. and Kritzman, M. (1999) Optimal portfolios in good times and bad. *Financial Analysts Journal* 55(3): 65–73.
- Engle, R. (2002) Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20(3): 339–350.
- Fengler, M. and Herwartz, H. (2008) Multivariate volatility models. In: W. Hardle, N. Hautsch and L. Overbeck (eds.) *Applied Quantitative Finance*. Berlin Heidelberg: Springer, pp. 313–326.
- Kritzman, M. and Li, Y. (2010) Skulls, financial turbulence, and risk management. *Financial Analysts Journal* 66(5): 30–41.

- Mahalanobis, P.C. (1927) Analysis of race-mixture in Bengal. *Journal of the Asiatic Society of Bengal* 23: 301–333.
- Mahalanobis, P.C. (1936) On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India* 2(1): 49–55.
- Markowitz, H. (1952) Portfolio selection. *Journal of Finance* 7(1): 77–91.

$$CS = \frac{1}{1-\rho^2} \times \left(1 - \frac{2\rho \frac{x}{\sigma_x} \frac{y}{\sigma_y}}{\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}} \right)$$

$$CS = \frac{1}{1-\rho^2} \times \left(1 - \frac{\rho z_x z_y}{\frac{1}{2}(z_x^2 + z_y^2)} \right)$$

APPENDIX

Correlation surprise mathematics for two assets

To build intuition around the correlation surprise methodology, we can investigate a simple two-asset example algebraically. To simplify the formulas, we will assume (without loss of generality) that the average value of x and average value of y are both zero.

$$CS = \frac{TI}{MS}$$

$$= \frac{(x \ y) \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix}}{(x \ y) \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix}}$$

Using the matrix identity that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

and multiplying out the matrices, we find that

$$CS = \frac{\sigma_x^2 \sigma_y^2}{(1-\rho^2) \sigma_x^2 \sigma_y^2} \times \frac{x^2 \sigma_y^2 + y^2 \sigma_x^2 - 2xy\rho\sigma_x\sigma_y}{x^2 \sigma_y^2 + y^2 \sigma_x^2}$$

$$CS = \frac{1}{1-\rho^2} \times \left(1 - \frac{2xy\rho\sigma_x\sigma_y}{x^2 \sigma_y^2 + y^2 \sigma_x^2} \right)$$

This formula highlights that correlation surprise is a function of the assets' correlation and their z -scores, which are volatility-normalized. All units expressing magnitude cancel out in the formula: the units of z -scores in the numerator equal the units of z -scores in the denominator. Hence, the correlation surprise score contains only information about the direction of co-movement, analogous to a compass or radial coordinate.

For this algebraic example we can determine that the minimum correlation surprise for a given correlation ρ is $(1-|\rho|)/(1-\rho^2)$ and the maximum is $(1+|\rho|)/(1-\rho^2)$. If correlation is zero, the minimum and maximum will both equal one. This makes sense, because if we do not expect any structured relationship between the two variables, no observed pattern of co-movement will be any more or less unusual than another. However, as the steady-state correlation intensifies, we will be increasingly less surprised by co-movement that is aligned with the steady-state correlation (we will see a CS score less than one), and increasingly more surprised by co-movement that diverges from the steady-state correlation (we will see a CS score greater than one).



This work is licensed under a Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>

Disclaimer: The views expressed in this article are the views solely of the authors and do not necessarily represent the views of State Street Corporation.